Super-rays based low rank approximation for light fields compression

Elian Dib, Mikaël Le Pendu, Xiaoran Jiang, Christine Guillemot

Super-rays based low rank approximation for light fields compression

Elian Dib*, Mikael Le Pendu†, Xiaoran Jiang *, Christine Guillemot*

*INRIA Rennes Bretagne Atlantique
Campus de Beaulieu
Rennes, Ille-et-Vilaine, 35 042, France
christine.guillemot@inria.fr

†Trinity College Dublin
College Green
Dublin, Ireland
lependum@scss.tcd.ie

Abstract

We describe a local low rank approximation method based on super-rays for light field compression. Super-rays can be seen as a set of super-pixels that are coherent across all light field views. A super-ray based disparity estimation method is proposed using a low rank prior, in order to be able to align all the super-pixels forming each super-ray. A dedicated super-ray construction method is described that constrains the super-pixels forming a given super-ray to be all of the same shape and size, dealing with occlusions. This constraint is needed so that the super-rays can be used as a support of angular dimensionality reduction based on low rank matrix approximation. A low rank matrix approximation is then computed on the disparity compensated super-rays using a singular value decomposition (SVD). A coding algorithm is then described for the different components of the resulting low rank approximation. Experimental results show performance gains compared with two reference light field coding schemes (HEVC-based scheme and JPEG-Pleno VM 1.1).

1 Introduction

Light field imaging has emerged as a promising technology for a number of applications going from photo-realistic image-based rendering to a variety of computer vision applications such as e.g. 3D modeling, object detection, classification, recognition. However, light fields represent very large volumes of high dimensional data, hence the need for designing efficient compression algorithms. In this paper, we focus on the problem of compression of dense light fields, as those captured by plenoptic cameras, which represent very large volumes of highly redundant data. While a number of methods have already been published in the literature aiming at adapting standardized solutions (in particular HEVC) to light field data as in [1–3], here we focus on the problem of reducing the angular dimension of light fields with a low rank approximation method. An homography-based low-rank approximation method called HLRA has been shown to give very good light fields compression performances in [4]. However, the low rank approximation in [4] is done globally on the entire light field.

In this paper, we explore the use of super-ray based local low rank approximation models for further improving the performance of the light field compression algorithm. The super-rays are used here as a way to better expose redundancy across the different views compared to a global homography-based alignment as done in [4], and in turn to further reduce the angular light field dimension. The concept of super-ray has been

This work has been supported by the EU H2020 Research and Innovation Programme under grant agreement No 694122 (ERC advanced grant CLIM).

initially introduced in [5] as an extension of super-pixels to address the computational complexity issue in light field image processing tasks. Super-rays can be seen as the clustering of rays of the light field within and across views, hence corresponding to the same set of 3D points of the imaged scene.

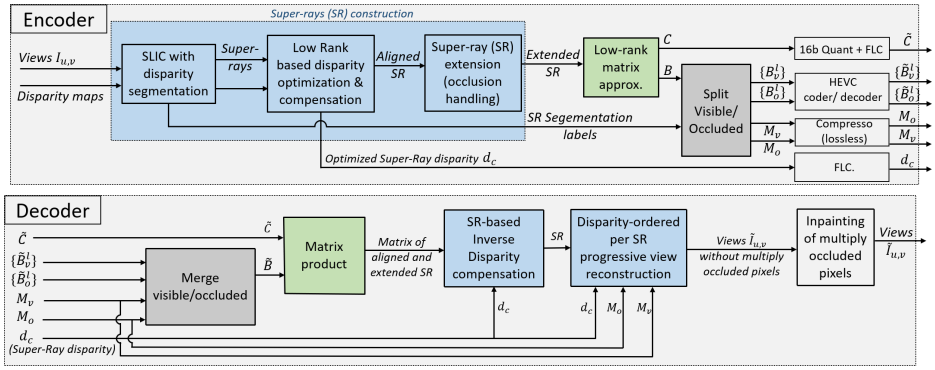

Figure 1: Overview of the super-ray based low rank approximation and coding chain (encoder), and Light field reconstruction steps (decoder).

In this paper, a dedicated super-ray construction method is first proposed so that super-rays can be used as the support of the low rank approximation. The central view is taken as a reference view on which super-pixels are constructed using the SLIC algorithm [6]. The super-rays which are constructed must group super-pixels which are consistent across the views while being constrained to be of same shape and size. The corresponding super-pixels in all views are found thanks to disparity compensation. In order to do so, a novel method is proposed to estimate the disparity for each super-ray using a low rank prior, so that the super-rays are constructed to yield the lowest approximation error for a given rank.

Due to occlusions, individual super-pixels may contain less pixels than the complete area captured in the corresponding super-ray (gathering information from super-pixels of all the views). The super-pixels are thus extended with the occluding pixels to satisfy the shape and size constraint. A low rank approximation is then computed for the set of extended super-rays that are stacked in a vectorized form in different columns of a matrix $\mathbf{X}$. The rank r approximation of the set of extended super-rays is expressed as a product of a matrix $\mathbf{B}$, containing r columns corresponding to basis (or eigen) images in a vectorized form, with a matrix $\mathbf{C}$ containing weighting coefficients. The matrix $\mathbf{B}$ is re-arranged into two sets of eigen images, respectively corresponding to the light rays visible and occluded in the central view. An image sequence formed with both those sets of 'occluded' and 'visible' eigen images is encoded using HEVC.

The method is also compared against one HEVC-based scheme which first forms a pseudo video sequence from all the views and then code this sequence with HEVC-inter coding [7], and with the verification model of JPEG-Pleno (VM 1.1) [8]. Experimental results show that a rate saving of up to -58.19% and -54.33% can be achieved compared with [7] and [8].

2 Related Work

This section does not give an exhaustive review of all light field compression techniques proposed so far but rather focuses on the most recent work on the topic. Existing solutions can be broadly classified into two categories: approaches directly compressing the lenslet images or approaches coding the views extracted from the raw data. Methods proposed for compressing the lenslet images mostly extend HEVC intra coding modes by adding new prediction modes to exploit similarity between lenslet images [1], [2] . In [9], a different approach is proposed using a depth based segmentation of the light field into 4D spatio angular blocks with prediction followed by JPEG2000.

Solutions extracting and coding views, code the views as pseudo video sequences using HEVC [7], [10], or the latest JEM coder [11], or extend HEVC to multi-view coding [12]. Methods using view synthesis have also been proposed in [13] using convolutional neural networks to synthesize all the views from the corner views, or in [14] where the authors use a linear approximation computed with Matching Pursuit for disparity based view prediction.

The most closely related work to the proposed method is the approach proposed in [15], in the sense that local models exploiting disparity are considered in the 4D rays space for compression. However, the models in [15] are 4D steered Gaussian Mixture Models (GMM), referred to as steered mixture of experts, while the proposed local models are based on low rank priors. The second most related work is the method in [4] introducing a global homography-based low rank approximation approach for light field compression. However, the proposed approach differs from the above approach, and extends it, by proposing a coding scheme based on local low rank models using a novel disparity estimator also exploiting low rank priors.

3 Notations and scheme overview

Let $\mathbf{L}(u, v, x, y)$ be a light ray of the light field $\mathbf{L}$, and (u, v, x, y) its coordinates using the two plane parametrisation, where (u, v) and (x, y) are the angular (view) and spatial (pixel) coordinates respectively. A view with angular coordinates (u, v) is denoted $\mathbf{I}_{u,v}$. The overall scheme is depicted in Fig.1 and comprises the following steps:

- Super-rays construction with the help of the disparity maps of all views.
- Low-rank based disparity optimization and compensation of each super-ray.
- Super-rays extension so that all super-pixels forming a SR are of same size.
- Low rank approximation of the set of extended super-rays.
- Splitting the eigen images into a first set $\{\mathbf{B}_v^l\}_{l\in[\![1,r]\!]}$ (with r the approximation rank) corresponding to the light rays visible in the central view, and a second set $\{\mathbf{B}_o^l\}$, very sparse, and containing eigen values of super-rays extensions.
- Encoding the information to be transmitted to the decoder, i.e., the two sets of eigen images $\{\mathbf{B}_v^l\}$ and $\{\mathbf{B}_o^l\}$, the matrix of coefficients $\mathbf{C}$, and side information needed at the decoder to find the super-rays and perform the inverse alignment.

These different steps are detailed below.

4 Super-ray construction with low rank based disparity optimization

Here, we present a super-ray construction method which constrains the projections of a super-ray on each view to be of same shape and size.

4.1 Light field over-segmentation

A k-means clustering method is applied on the entire light field in order to group pixels of all the views which are in local regions and which are similar in colour (in CIELAB colourspace) and in disparity. This clustering, extending the concept of super-pixels, forms so-called super-rays since it groups pixels (and super-pixels) across views which are assumed to correspond to the same set of 3D points in space. This leads to a segmentation of all the views into super-pixels which are coherent across the views. The clustering into super-rays is performed using the disparity estimation method of [16]. We then assume that all pixels belonging to a given super-ray have the same disparities referred to as the *centroid disparities*.

To compute the low rank approximation, the super-pixels in each super-ray need to be aligned by disparity compensation. We show in the experimental section that disparity estimated independently of the low rank constraint, e.g. as in [16], is not optimal for our low rank based coding scheme. We hence propose in the next section a novel disparity optimization method for each super-ray based on a low rank prior.

4.2 Disparity estimation of super-rays using a low rank prior

We use a low rank prior on each super-ray as a means to locally approximate the light field and jointly perform disparity compensation so as to align similar pixels from different views and render the approximation even more efficient. For a given super-ray, let h^d be the set of disparity compensation operators h_i^d for each super-pixel at a view (u_i, v_i) forming the super-ray. The disparity compensation h_i^d is defined as $h_i^d : (x, y) \mapsto (x+d_x(u_i-u_c), y+d_y(v_i-v_c))$, and d is the vector $\begin{pmatrix} d_x & d_y \end{pmatrix}^\top$ to estimate.

The disparity estimation algorithm proceeds as follows. The horizontal and vertical disparity values d_x and d_y are first initialized to "0". Note that even after disparity compensation, the aligned super-pixels forming a super-ray may have different shapes and sizes, in particular due to occlusions, which would make the low rank approximation impractical. To cope with this difficulty, super-pixels with missing information are filled using occluding pixels. The set of aligned and completed super-pixels is called *extended* super-ray. All super-pixels forming a given extended super-ray $\mathcal{SR}$ are stored, in a vectorized form, in a matrix

$$\mathcal{SR} = [\mathcal{SP}_1 \mid \mathcal{SP}_2 \mid ...], \tag{1}$$

where $\mathcal{SP}_i$ is the vectorized version of the super-pixel from view (u_i, v_i).

For each extended super-ray, we then solve

$$\text{argmin}_{d,\mathbf{M}} \left\| \mathcal{SR} \circ h^d - \mathbf{M} \right\|_F^2 \text{ s.t. } \text{rank}(\mathbf{M}) = r, \tag{2}$$

where $\mathbf{M}$ is a rank r matrix and $\|.\|_F$ is the Frobenius norm. The notation "$\circ$" stands for the application of the disparity compensations to align all the super-pixels forming a super-ray.

We iteratively solve the problem by successively finding $\mathbf{M}$ while fixing d, and updating d while fixing $\mathbf{M}$, at each iteration, until convergence. For a fixed d, the matrix $\mathbf{M}$ of rank r is obtained as $\mathbf{M} = \mathbf{U}\mathbf{\Sigma}_r\mathbf{V^T}$ where $\mathbf{U}\mathbf{\Sigma}\mathbf{V^T}$ is the singular value decomposition (SVD) of $\mathcal{SR} \circ h^d$ and $\mathbf{\Sigma}_r$ contains only the r largest singular values of $\mathbf{\Sigma}$. The set disparity d is updated, by solving (2) for the rank r matrix $\mathbf{M}$ fixed. Minimizing (2) is not trivial due to the non linearity of the term $\mathcal{SR} \circ h^d$.

Assuming the change Δd at each iteration to be small, we can approximate this update by local linearity as

$$\mathcal{SP}_i \circ h_i^{d+\Delta d} \approx \mathcal{SP}_i \circ h_i^d + \frac{\partial[\mathcal{SP}_i \circ h_i^d]}{\partial d_x}\Delta d_x + \frac{\partial[\mathcal{SP}_i \circ h_i^d]}{\partial d_y}\Delta d_y \tag{3}$$

$$= \mathcal{SP}_i \circ h_i^d + \frac{\partial[\mathcal{SP}_i \circ h_i^d]}{\partial x}(u_i - u_c)\Delta d_x + \frac{\partial[\mathcal{SP}_i \circ h_i^d]}{\partial y}(v_i - v_c)\Delta d_y \tag{4}$$

Given this approximation, d is updated with the increment $\widehat{\Delta d}$ obtained by solving

$$\widehat{\Delta d} = \text{argmin}_{\Delta d} \sum_i \|\mathbf{R}_i - \mathbf{J}_i \Delta d\|_F^2 , \tag{5}$$

with, $\mathbf{R}_i = \mathbf{M}_i - \mathcal{SP}_i \circ h_i^d$ and $\mathbf{J}_i = \begin{pmatrix}(u_i - u_c)\frac{\partial[\mathcal{SP}_i \circ h_i^d]}{\partial x} & (v_i - v_c)\frac{\partial[\mathcal{SP}_i \circ h_i^d]}{\partial y}\end{pmatrix}$.

This problem has the following analytical solution:

$$\widehat{\Delta d} = \left(\sum_i J_i^\top J_i\right)^{-1} \sum_i J_i^\top R_i \tag{6}$$

Equation 2 minimizes the approximation error of the matrix formed by the aligned views. This approximation error corresponds to the PSNR that we call PSNRin in the experimental section. We however show in the sequel that significantly better results are obtained by taking into account the inverse warping (i.e. the corresponding interpolation errors) in the minimization. At each iteration, for a given d we evaluate $\left\|\mathcal{SR} - (h^d)^{-1} \circ \mathbf{M}\right\|_F^2$ to eventually keep the disparity values d minimizing the super-ray reconstruction error, corresponding to a PSNR that we call PSNRout.

5 Low Rank Approximation

All the matrices corresponding to the different extended and aligned super-rays $\mathcal{SR}_k$ are stacked in a matrix $\mathbf{X}$ of dimension $\mathbb{R}^{m \times n}$, where n is the number of views and m is the number of pixels per view of the concatenated extended super-rays. In other words, the matrix $\mathbf{X}$ contains all the pixels of all the views that are locally aligned super-ray per super-ray.

The matrix $\mathbf{X}$ is factorized into a product of a low-rank matrix $\mathbf{B} \in \mathbb{R}^{m \times r}$ and a coefficient matrix $\mathbf{C} \in \mathbb{R}^{r \times n}$, with $r \leq n$, as

$$\text{argmin}_{\mathbf{B},\mathbf{C}} \|\mathbf{X} - \mathbf{B}\mathbf{C}\|_F^2 , \tag{7}$$

where $\|.\|_F$ is the Frobenius norm. Optimal factorization is obtained from the singular value decomposition (SVD) of X into $U\Sigma V^\top$. Assuming the singular values in Σ are in decreasing order, we take B as the r first columns of $U\Sigma$, and C as the r first rows of $V^\top$. The set of super-rays $\{\mathcal{SR}\}$ forming the entire light field can thus be approximated by a linear combination of columns of $\mathbf{B}$, therefore significantly reducing the amount of data at the cost however of an approximation error.

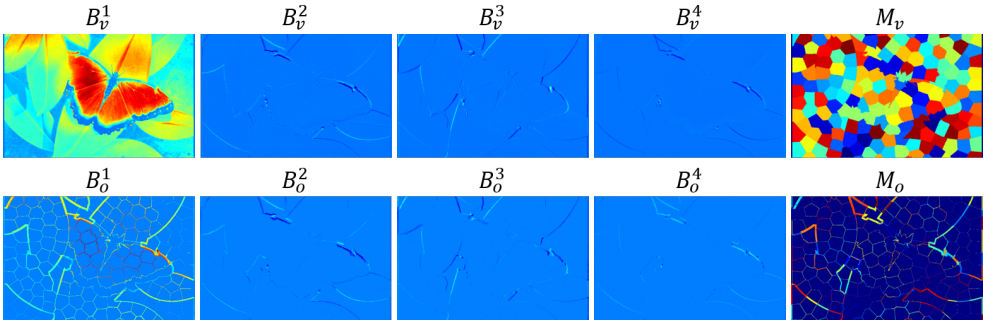

Figure 2: First 4 eigen-images (shown in false colors) and segmentation maps for the visible ($\mathbf{B}_v^l$, $\mathbf{M}_v$) and occluded ($\mathbf{B}_o^l$, $\mathbf{M}_o$) sets of pixels ("butterfly" synthetic light field [17]).

6 Compression Scheme

The above low rank approximation leads to a matrix $\mathbf{B}$ containing in its columns eigen images of the extended views (set of extended super-pixels) and to a matrix $\mathbf{C}$ containing coefficients. For the purpose of encoding the matrix $\mathbf{B}$ using HEVC, the data in each column l is re-arranged into two eigen images $\mathbf{B}_v^l$ and $\mathbf{B}_o^l$. $\mathbf{B}_v^l$ contains the entries derived from the portions of the extended super-rays which are visible in the central view. $\mathbf{B}_o^l$ contains the entries corresponding to areas occluded in the central view. Two segmentation maps $\mathbf{M}_v$ and $\mathbf{M}_o$ are also computed so that each region in $\mathbf{M}_v$ (respectively $\mathbf{M}_o$) is labelled with the index of the super-ray that will be reconstructed from the collocated region in the $\mathbf{B}_v^l$ images (respectively $\mathbf{B}_o^l$). Examples of eigen images $\mathbf{B}_v^l$ and $\mathbf{B}_o^l$ with their associated maps $\mathbf{M}_v$ and $\mathbf{M}_o$ are illustrated in Fig.2.

The data in the images $\mathbf{B}_v^l$ and $\mathbf{B}_o^l$ and the matrix C are quantized on 16 bits. The quantized entries of the $\mathbf{C}$ matrix are directly transmitted (e.g. coded using a fixed length code) since the corresponding cost is quite negligible. As shown in Fig.2, the first 4 eigen images $\mathbf{B}_v^l$ and $\mathbf{B}_o^l$, show some correlation. Hence to code them efficiently using HEVC-Inter, the empty regions of the $\mathbf{B}_o^l$ images are completed with the collocated regions of the corresponding $\mathbf{B}_v^l$ images. The $\mathbf{B}_v^l$ and $\mathbf{B}_o^l$ images are then interleaved as $\mathbf{B}_v^1$, $\mathbf{B}_o^1$, ..., $\mathbf{B}_v^4$, $\mathbf{B}_o^4$, and coded using HEVC-Inter with a GOP size of 2. The remaining columns of these matrices being less correlated are coded using HEVC-Intra.

It is also necessary to transmit as side information the two centroid disparity values per super-ray as well as the two segmentation maps $\mathbf{M}_v$ and $\mathbf{M}_o$. The two segmentation maps are compressed without loss using the compresso method described in [18]. The disparity information (two values d_x, d_y per super-ray) is encoded using a fixed length code on 32 bits, which is quite negligible given the small number of values to be transmitted (in the experiments we considered a partition of the light fields into 200 super-rays for natural light fields and 240 for the synthetic ones).

7 Light field reconstruction

On the receiver side, the decoded sequence of images $\tilde{\mathbf{B}}_v^l$ and $\tilde{\mathbf{B}}_o^l$ need to be merged to form the decoded low rank matrix $\tilde{\mathbf{B}}$ which contain the eigen images of all the aligned and extended super-rays. This matrix $\tilde{\mathbf{B}}$ is multiplied by the matrix $\tilde{\mathbf{C}}$ to obtain the

matrix $\tilde{\mathbf{X}}$ of aligned and extended super-rays. An inverse disparity compensation is then performed per super-ray using its centroid disparities. Each view of the light field is then reconstructed by progressively mapping each super-ray on each view starting from the super-rays having the smallest disparity, i.e. starting from background pixels. Pixels having a higher disparity over-write previously rendered pixels at the same position with a lower disparity.

Depending on the complexity of the scene, on the number of objects and depth layers, some parts in the light field may be occluded by several objects. The corresponding pixels are referred to here as *multiply-occluded* pixels. One sets of eigen images $\{\mathbf{B}_o^l\}_l$ is not sufficient to represent *multiply-occluded* pixels. One alternative would be to transmit additional sets $\{\mathbf{B}_o^l\}_l$ to represent these pixels. However, the number of such pixels is quite limited and the corresponding matrix would be very sparse. Instead, we inpaint the corresponding pixels in the reconstructed light field $\tilde{\mathbf{L}}$ using a low rank completion matrix. The low rank matrix completion problem is posed as the search for the minimum nuclear norm matrix $\hat{\mathbf{L}}$ with entries equal to those of the matrix $\tilde{\mathbf{L}}$ for the known elements of $\tilde{\mathbf{L}}$. The problem is mathematically formulated as

$$\min_{\hat{\mathbf{L}}} ||\hat{\mathbf{L}}||_* \text{ s.t. } \forall (i,j) \in \Omega, \hat{\mathbf{L}}_{ij} = \tilde{\mathbf{L}}_{ij}, \tag{8}$$

where Ω is the set of indices of the known elements in $\tilde{\mathbf{L}}$, and $||.||_*$ is the nuclear norm (convex approximation of the rank). This minimization is solved using the Inexact ALM (IALM) technique [19].

8 Experimental Results

The performances of the proposed disparity estimation and light field compression methods have been evaluated for the luminance component of light fields shown in Fig.4 and coming from : 1/- INRIA dataset [20] which contains LFs captured by a second generation Lytro Illum camera from which we use the 9×9 central sub-aperture images of 625×434 pixels each; 2/- the HCI dataset [17] synthetic LFs 9×9 views of 768×768 pixels; 3/- the ICME 2016 Grand Challenge dataset [21] from which we take the 9×9 central sub-aperture images. The Lytro LFs have been decoded using the Matlab Light Field Toolbox v0.4 [22] with gamma correction. The methods involving a low rank prior (HLRA and our method Local Low Rank Approximation (LLRA)) were evaluated for varying target rank values $r = \{1, 3, 5, 15, 30, 60\}$ and HEVC quality parameter $QP \in \{2, 6, 14, 20, 26, 38\}$. We retain the (r,QP) pairs corresponding to the points on the convex enveloppe of the rate distortion graph. For HEVC Lozenge $QP = \{10, 14, 17, 20, 23, 26\}$. The disparity maps used for LLRA and JPEG Pleno were generated using [16]. For JPEG Pleno we adapted the existing $Fountain\_Vincent\_2$ configuration files for use on 9×9 LFs and used it for every LF.

8.1 Performance analysis of the disparity estimation using a low rank prior

We first assess in Fig.3 the performance of the proposed super-ray disparity estimation method based on a low rank prior against an exhaustive search, and disparity

computed independently of the low rank constraint (in the experiments we used disparity values computed with [16]). In the exhaustive search approach, the optimal disparity values for each super-ray are obtained by evaluating the low rank approximation error for a wide range of integer, and then fraction-pixel refined, disparity values. We observe that using a low rank prior improves significantly the performance compared to using input disparity maps computed with no rank constraint. Moreover, our method compares favorably with the exhaustive search. We also observe that aiming at maximizing the PSNR after inverse warping (*PSNRout*) during the SR disparity optimization rather than maximizing the PSNR of the low rank approximation of the aligned SR (*PSNRin*) allows us to mitigate the information loss due to bi-cubic interpolation in forward and inverse warping. In the sequel, we use the low-rank based and exhaustive search disparity estimation with maximization of outter PSNR.

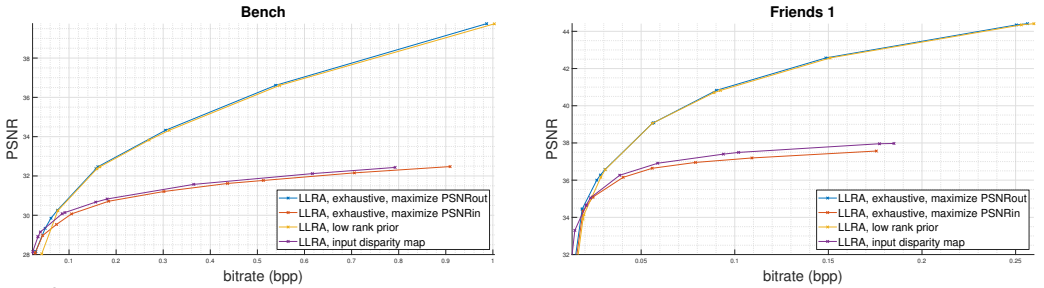

Figure 3: PSNR-Rate performance of the low rank based disparity estimation compared with an exhaustive search and the use of a disparity map computed with no rank constraint.

8.2 Performance analysis of the complete scheme

Tab. 1 shows the Bjontegaard rate savings obtained for the HLRA method [4] and two variants of our method with respect to the HEVC-Lozenge [7] and JPEG Pleno Verification Model [8] anchors and we consistently observe rate savings on natural light fields compared with the two anchors. In particular, for real light fields, we observe comparable performance by using the low rank prior to compute the centroid disparities or by performing an exhaustive search. However, for synthetic light fields, the lower performance is due to the initialization of the gradient descent ($(d_x, d_y) = (0, 0)$) that is not appropriate for larger baselines. By initializing the centroid disparities of each super-ray to the median disparity of the pixels of a disparity map [16], we observe in Fig.5 a better performance for LLRA for light fields with larger baselines such as StillLife and Buddha. For these light fields, LLRA can outperform HLRA despite the extra signalling cost due to the segmentation in super-rays.

9 Conclusion

This paper has presented a compression scheme for light fields using super-ray based local low rank models. A novel method for disparity estimation and compensation was proposed so that the super-rays are constructed to yield the lowest approximation error for a given rank. The proposed representation is actually based on two low rank models, one for the central view pixels that are visible in all views and one for the occlusions. We could observe consistent bitrate savings on all natural light fields

Figure 4: Test light fields. Top row (first two): HCI (Buddha, StillLife) [17]. Top row (last three): INRIA Dataset (Bench, Fruits, Toys) [20]; Bottom row: ICME Dataset (Fountain_Vincent_2, Friends_1, StonePillars, Vespa) [21].

Method	HLRA		LLRA (exhaustive disparity search)		LLRA (low rank prior disparity)	
Anchor	HEVC Lozenge	JPEG Pleno	HEVC Lozenge	JPEG Pleno	HEVC Lozenge	JPEG Pleno
F_V_2	-40.93%	-51.58%	-14.41%	-8.28%	-11.91%	-5.65%
Friends_1	-65.62%	-73.52%	-57.44%	-43.07%	-58.19%	-42.87%
S_P_I	-67.84%	-72.45%	-54.95%	-54.33%	-50.24%	-50.73%
Vespa	-49.57%	-61.21%	-37.85%	-12.61%	-36.54%	-5.09%
Bench	-67.37%	-64.25%	-46.56%	-25.56%	-45.65%	-25.16%
Fruits	-55.15%	-59.19%	-26.43%	-12.54%	-27.05%	-12.74%
Toys	-58.98%	-81.05%	-41.32%	-48.96%	-42.09%	-48.90%

Table 1: Bjontegaard rate savings for HLRA and LLRA compression schemes with respect to HEVC-Lozenge and JPEG Pleno VM anchors.

compared to reference methods (HEVC and JPEG-Pleno). However, compared to global homographies (HLRA), the segmentation in super-rays implies sending extra side information that can be penalizing at low rate, for light fields with small baselines. However, for larger baselines, the performances benefit from the better alignment and approximation of the local model.

[1] C. Conti, P. Nunes, and L. D. Soares, "HEVC-based light field image coding with bi-predicted self-similarity compensation," in *IEEE Int. Conf. Multimed. Expo Workshops (ICMEW)*, Jul. 2016.

[2] R. Monteiro, L. Lucas, C. Conti, P. Nunes, N. Rodrigues, S. Faria, C. Pagliari, E. da Silva, and L. Soares, "Light field HEVC-based image coding using locally linear embedding and self-similarity compensated prediction," in *IEEE Int. Conf. Multimed. Expo Workshops (ICMEW)*, Jul. 2016.

[3] Y. Li, M. Sjöström, R. Olsson, and U. Jennehag, "Efficient intra prediction scheme for light field image compression," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 539–543.

[4] X. Jiang, M. Le Pendu, R. Farrugia, and C. Guillemot, "Light field compression with homography-based low-rank approximation," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 1132–1145, Oct. 2017.

[5] M. Hog, N. Sabater, and C. Guillemot, "Super-rays for efficient light field processing," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 1187–199, Oct. 2017.

[6] R. Achanta, A. Shaji, Kevin K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, 2012.

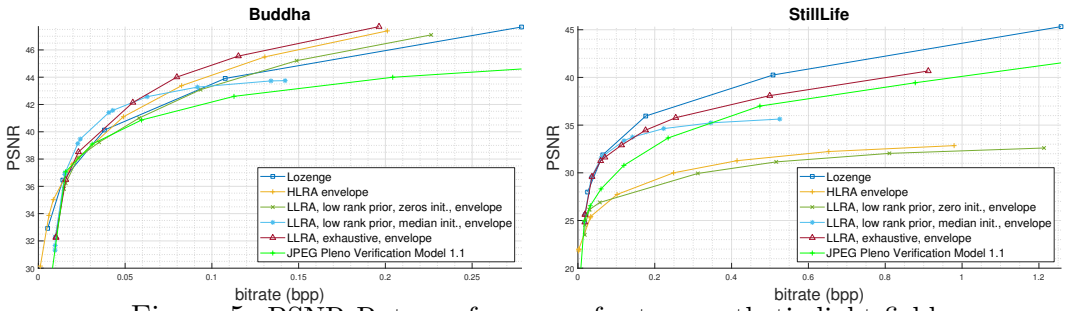

Figure 5: PSNR-Rate performance for two synthetic light fields.

[7] M. Rizkallah, T. Maugey, C. Yaacoub, and C. Guillemot, "Impact of light field compression on focus stack and extended focus images," in *European Signal Processing Conf. (EUSIPCO)*, Aug. 2016, pp. 898–902.

[8] ISO/IEC JTC 1/SC29/WG1 JPEG, "Jpeg pleno light field coding vm 1.1," Doc. N81052, 2018.

[9] I. Tabus, P. Helin, and P. Astola, "Lossy compression of lenslet images from plenoptic cameras combining sparse predictive coding and jpeg 2000," in *IEEE Int. Conf. Image Process. (ICIP)*. IEEE, 2017, pp. 4567–4571.

[10] D. Liu, L. Wang, L. Li, Z. Xiong, F. Wu, and W. Zeng, "Pseudo-sequence-based light field image compression," in *IEEE Int. Conf. Multimed. Expo Workshops (ICMEW)*, Jul. 2016.

[11] C. Jia, Y. Yang, X. Zhangy, X. Zhang, S. Wangx, S. Wang, and S. Ma, "Optimized inter-view prediction based light field image compression with adaptive reconstruction," in *IEEE Int. Conf. on Image Processing, ICIP*, 2017.

[12] W. Ahmad, R. Olsson, and M. Sjostrom, "Interpreting plenoptic images as multiview sequences for improved compression," in *IEEE Int. Conf. Image Process. (ICIP)*, 2017.

[13] X. Jiang, M. Le Pendu, and C. Guillemot, "Light fields compression using depth image based view synthesis," in *Hot3D workshop held jointly with IEEE Int. Conf. Multimed. Expo, ICME*, Jul. 2017.

[14] S. Zhao and Z. Chen, "Light field image coding via linear approximation prior," in *IEEE Int. Conf. on Image Processing, ICIP*, 2017.

[15] R. Verhack, T. Sikora, L. Lange, R. Jongebloed, G. Van Wallendael, and P. Lambert, "Steered mixture-of-experts for light field coding, depth estimation, and processing," in *IEEE Int. Conf. Multimed. Expo Workshops (ICMEW)*, 2017, pp. 1183–1188.

[16] X. Jiang, M. Le Pendu, and C. Guillemot, "Depth estimation with occlusion handling from a sparse set of light field views," in *IEEE Int. Conf. Image Process. (ICIP)*, 2018.

[17] S. Wanner, S. Meister, and B. Goldluecke, "Datasets and benchmarks for densely sampled 4D light fields," in *VMV Workshop*, 2013, pp. 225–226.

[18] B. Matejek, D. Haehn, F. Lekschas, M. Mitzenmacher, and H. Pfister, "Compresso: Efficient compression of segmentation data for connectomics," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2017.

[19] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," Tech. Rep., University of Illinois at Urbana-Champaign, 2009.

[20] "INRIA Lytro image dataset," `https://www.irisa.fr/temics/demos/lightField/LowRank2/datasets/datasets.html`.

[21] "ICME 2016 Grand Challenge dataset," `http://mmspg.epfl.ch/EPFL-light-field-image-dataset`.

[22] D.G. Dansereau, "Light Field Toolbox for Matlab," 2015.